

PCT

WORLD INTELLECTUAL PROPERTY ORGANIZATION
International Bureau

INTERNATIONAL APPLICATION PUBLISHED UNDER THE PATENT COOPERATION TREATY (PCT)

<p>(51) International Patent Classification ⁶ : G06K 9/20 // G06F 17/24</p>	<p>A1</p>	<p>(11) International Publication Number: WO 98/47098</p> <p>(43) International Publication Date: 22 October 1998 (22.10.98)</p>
<p>(21) International Application Number: PCT/SE98/00602</p> <p>(22) International Filing Date: 1 April 1998 (01.04.98)</p> <p>(30) Priority Data: 9701183-7 1 April 1997 (01.04.97) SE</p> <p>(71) Applicant (for all designated States except US): READSOFT AB [SE/SE]; Garnisionsgatan 25A, S-254 66 Helsingborg (SE).</p> <p>(72) Inventor; and (75) Inventor/Applicant (for US only): ANDERSSON, Jan [SE/SE]; Banckagatan 17, S-260 40 Viken (SE).</p> <p>(74) Agents: ONN, Thorsten et al.; AB Stockholms Patentbyrå, Zacco & Bruhn (publ), P.O. Box 23101, S-104 35 Stockholm (SE).</p>		<p>(81) Designated States: AL, AM, AT, AU, AZ, BA, BB, BG, BR, BY, CA, CH, CN, CU, CZ, DE, DK, EE, ES, FI, GB, GE, GH, GM, GW, HU, ID, IL, IS, JP, KE, KG, KP, KR, KZ, LC, LK, LR, LS, LT, LU, LV, MD, MG, MK, MN, MW, MX, NO, NZ, PL, PT, RO, RU, SD, SE, SG, SI, SK, SL, TJ, TM, TR, TT, UA, UG, US, UZ, VN, YU, ZW, ARIPO patent (GH, GM, KE, LS, MW, SD, SZ, UG, ZW), Eurasian patent (AM, AZ, BY, KG, KZ, MD, RU, TJ, TM), European patent (AT, BE, CH, CY, DE, DK, ES, FI, FR, GB, GR, IE, IT, LU, MC, NL, PT, SE), OAPI patent (BF, BJ, CF, CG, CI, CM, GA, GN, ML, MR, NE, SN, TD, TG).</p> <p>Published With international search report. In English translation (filed in Swedish).</p>
<p>(54) Title: METHOD AND ARRANGEMENT FOR AUTOMATIC DATA ACQUISITION OF FORMS</p> <p>(57) Abstract</p> <p>The invention relates to a method and an arrangement for the automatic data acquisition of paper-based information, whereby an unknown form (10) that is scanned into a computer is transformed into line elements (14, 15) for quicker identification. The identification of known forms (10) is also simplified by means of the line elements (14, 15).</p> <div data-bbox="727 1171 1380 1873"> <p>The diagram shows a scanned form (10) at the top left. An arrow points from the form to a set of line elements (14, 15) below it. The line elements are represented as a grid of horizontal and vertical lines. A coordinate system (x, y) is shown at the bottom left. Labels include 'ReadSoft AB' and '12' near the top left, and '14', '15', '16', '18' near the line elements and the form.</p> </div>		

BEST AVAILABLE COPY

FOR THE PURPOSES OF INFORMATION ONLY

Codes used to identify States party to the PCT on the front pages of pamphlets publishing international applications under the PCT.

AL	Albania	ES	Spain	LS	Lesotho	SI	Slovenia
AM	Armenia	FI	Finland	LT	Lithuania	SK	Slovakia
AT	Austria	FR	France	LU	Luxembourg	SN	Senegal
AU	Australia	GA	Gabon	LV	Latvia	SZ	Swaziland
AZ	Azerbaijan	GB	United Kingdom	MC	Monaco	TD	Chad
BA	Bosnia and Herzegovina	GE	Georgia	MD	Republic of Moldova	TG	Togo
BB	Barbados	GH	Ghana	MG	Madagascar	TJ	Tajikistan
BE	Belgium	GN	Guinea	MK	The former Yugoslav Republic of Macedonia	TM	Turkmenistan
BF	Burkina Faso	GR	Greece	ML	Mali	TR	Turkey
BG	Bulgaria	HU	Hungary	MN	Mongolia	TT	Trinidad and Tobago
BJ	Benin	IE	Ireland	MR	Mauritania	UA	Ukraine
BR	Brazil	IL	Israel	MW	Malawi	UG	Uganda
BY	Belarus	IS	Iceland	MX	Mexico	US	United States of America
CA	Canada	IT	Italy	NE	Niger	UZ	Uzbekistan
CF	Central African Republic	JP	Japan	NL	Netherlands	VN	Viet Nam
CG	Congo	KE	Kenya	NO	Norway	YU	Yugoslavia
CH	Switzerland	KG	Kyrgyzstan	NZ	New Zealand	ZW	Zimbabwe
CI	Côte d'Ivoire	KP	Democratic People's Republic of Korea	PL	Poland		
CM	Cameroon	KR	Republic of Korea	PT	Portugal		
CN	China	KZ	Kazakhstan	RO	Romania		
CU	Cuba	LC	Saint Lucia	RU	Russian Federation		
CZ	Czech Republic	LI	Liechtenstein	SD	Sudan		
DE	Germany	LK	Sri Lanka	SE	Sweden		
DK	Denmark	LR	Liberia	SG	Singapore		
EE	Estonia						

Title

Method and arrangement for automatic data acquisition of forms

Technical field

The present invention refers to a method and arrangement for the automatic data
5 acquisition, by means of a means for the same, of forms whose design and information is not
known in advance, by input into the said means together with storage of patterns of the same.

The prior art

It is a problem for companies, organisations and others to make good use of the
information found in different types of paper forms, documents, etc.

10 With new, modern technology, these items can be scanned with a scanner and
entered into a database via commercially available software programs. However, sorting,
identification and other checking routines must to a large extent still be performed manually via
the computer's display or screen.

For example, to store an invoice from one and the same company as one
15 specifically designed document with a logotype and other visual elements, it must be revised so
that its format is adapted to one that can be accepted by the software and then stored in a
database. This and other procedures must be repeated each time an invoice with a new design is
scanned with the software.

To identify an invoice from a company that is already registered, the whole invoice
20 is often searched, which is time consuming. Certain software programs can have search routines
that restrict the extent of this searching. It is, however, difficult to safeguard against blurred or
hand-written lines of text, etc.

A need therefore exists for all who handle invoices and other forms to quickly be
able to identify these and/or quickly be able to enter and store new patterns in their invoicing
25 system.

Patent US-A-4 933 979 describes traditional data acquisition from forms and
requires pre-defined templates/patterns with no self-learning (adaptive) ability.

Patent US-A-5 140 650 mentions data acquisition from forms with what is known as
"Form out" technology to cover-over the original document and only retain the parts that are
30 "filled-in". This data acquisition is often combined with data acquisition according to US-A-4
933 979. The patent does not have any adaptive function for data acquisition of unknown forms.

Another patent, US-A-5 293 429, concerns the classification of documents with the help of lines on the documents and does not directly concern data acquisition or any adaptive function for this. US-A-5 293 429 does not ensure the identification of lines with object areas (areas with text) and a "RCG-value" (ReCoGnition, a number that uniquely identifies a document).

None of the said patents generates a form map for a form unknown to the system according to the patent and stores the map in real time in a form database for recognition at the next opportunity for identification. For the inventions according to these patents, the unknown form must therefore be stored later by other means.

Summary of the invention

One of the objectives of the present invention is to solve the problems named above as well as others during what is known as automatic data acquisition (interpretation) in connection with the handling of paper-based information.

The present invention concerns a system (method and arrangement) for the automatic data acquisition of forms where the system has no prior knowledge of what the form looks like or where on the form the information is to be found. In this way, templates of forms do not have to be defined in advance, but are instead registered as they are submitted to the system, i.e., in real time.

To accomplish the above objectives, the present invention specifies a method and arrangement for the automatic data acquisition, by means of a means for the same, of forms whose design and information is not known in advance, by input into the said means together with storage of patterns of the same. The method is adaptive, by which it includes learning and registering of forms as patterns without filled-in text, and by it also including the following steps for accomplishing the adaptive registration:

generation of a form map based on the previously unknown form's design for identifying information contained on the form;

searching and comparing the form map with stored, registered maps in a means for storing form maps;

storing generated form maps in the storage means when they do not coincide with a stored map according to pre-determined limits for agreement;

indication of agreement according to the limits for agreement when agreement is found; and

continued data acquisition for identifying of the information content of the form.

According to one embodiment of the present invention, the form map can consist of an object area list with objects contained in the form whereby the object comprises colours and/or wholly or partly of text.

5 In an alternative embodiment, the form map constitutes a line map comprising objects in the form of coloured lines from the form.

Horizontal lines in the line map are used to produce a horizontal key by dividing the form into a pre-determined number of horizontal segments along the y-axis in a cartographic system of co-ordinates, whereby each segment is equivalent to a position in the horizontal key.

10 Vertical lines in the line map are used to produce a vertical key by dividing the form into a pre-determined number of vertical segments along the x-axis in a cartographic system of co-ordinates, whereby each segment is equivalent to a position in the vertical key.

At least one line element that is included in a segment is marked in the equivalent key position, and segments that lack line elements remain unmarked in the equivalent key position.

15 A horizontal key and/or a vertical key constitute a line key in the line map, whereby during the said searching, the line key generated is compared with line keys stored in the means for verifying agreement.

The line keys are sorted in the storage means according to the number of markings.

20 The object's horizontal position in the object area list is used to generate a horizontal key by dividing the form into a pre-determined number of horizontal segments along the y-axis in a cartographic system of co-ordinates, whereby each segment is equivalent to a position in the horizontal key.

25 The object's vertical position in the object area list is used to produce a vertical key by dividing the form into a pre-determined number of vertical segments along the x-axis in a cartographic system of co-ordinates, whereby each segment is equivalent to a position in the vertical key.

At least one object that is included in a segment is marked in the equivalent key position, and segments that lack objects remain unmarked in the equivalent key position.

30 A horizontal key and/or a vertical key constitute an object key in the object area list, whereby during the said searching, the object key generated is compared with object keys stored in the means for verifying agreement.

The object keys are preferably sorted in the storage means according to the number of markings.

Searching results in a pre-defined number of requested probable candidates for the currently searched form.

5 If needed, an operator can support manually the whole or parts of the adaptive registration or identification of the new form or registered forms respectively if several alternative candidates are found as probabilities according to a factor of merit.

Finally, the identity of the form is confirmed by the data acquisition of a RCG-value.

10 Furthermore, the present invention specifies a arrangement for performing the above method.

The arrangement carries out automated data acquisition, by means of a means for the same, of forms whose design and information is not known in advance, by input into the said arrangement together with storage of patterns of the same. It learns adaptively and registers the design of the form, and includes a computer with the following means for carrying out the adaptive registration:

means for generation of a form map based on the previously unknown form's design for identifying information contained on the form;

20 means for searching and comparing the form map with stored, registered maps in a means for storing form maps;

means for storing generated form maps in the storage means when they do not coincide with a stored map according to pre-determined limits for agreement;

means for indicating agreement according to the limits for agreement when agreement is found; and

25 means for identification and continued data acquisition of the information content of the form.

In addition, the arrangement can include or constitute that specified according to the above method of the present invention, which is further illustrated in the accompanying non-independent claims for the arrangement.

30 **Brief description of the drawings**

Further reference to the enclosed figures and associated text will give a clearer understanding of the present invention.

Fig. 1 shows schematically how a line pattern is accomplished from a scanned-in invoice.

Fig. 2 shows schematically a flow-path for scanning, identifying, interpreting and validating a form according to the present invention.

5

Detailed description of preferred embodiments

In the continued description of the present invention, the forms are presented as invoices. However, the invention is not limited to invoice forms but also covers general documents containing text, figures, etc. as forms. Invoices are used here as an example of a form to exemplify the invention.

10

Fig. 1 illustrates schematically one part of invoice 10 that is scanned in a computer and that is shown on the display. As is evident from invoice 10, it is unclear or blurred after the scanning or input.

Invoice 10 consists partly of a logotype 12 and the vertical 14 and horizontal line 15 elements.

15

Note that even the logotype contains long black or varying degrees of shaded coloured line elements 16 that have been partly registered in a line map 18 according to the present invention, and that give an idea of what the original logotype 12 looked like, which simplifies identification when the invoice is an object to be identified as being as registered in a form map database. Coloured lines also include grey scales of black.

20

The form map that in this case constitutes line map 18 has been filtered from other objects 19, such as whole or parts of text objects or coloured objects, plus even the said line elements that include colour, which cannot be reproduced here, but that can be included as many coloured fields on a form 10.

25

An invoice 10 that is prepared according to the present invention, hereafter designated EH (Eyes & Hands), must be identified at an early stage. For successful identification, EH must have on a previous occasion, learned what the current invoice 10 looks like, which in simple terms means that information about the invoice is available in the form database in EH.

30

By necessity, the identification must be quick and be able to be made in a database that holds a very large number of invoice identities 18. It is not uncommon for databases to contain more than 10,000 identities 18.

The method and arrangement that EH uses does not require that an invoice is always put through a scanner in exactly the same way, i.e. the information on the invoice can vary somewhat in the x and y axes within a pre-determined measurement or threshold value. Fig. 1 shows a schematic cartographic system of co-ordinates.

5 In the present invention, (EH) comprises that in one embodiment, EH searches for all vertical 14 and horizontal line 15 elements of a pre-determined length on the invoice. Lines 14, 15 do not need to be free-standing and isolated, but can, for example, be part of a larger logotype text 12, such as ReadSoft AB in Fig. 1. The logotype 12 is represented as the line element 16 in line map 18.

10 The horizontal lines 15 and the vertical lines 14 constitute the basis for generating a horizontal key (h-key) and a vertical key (v-key) respectively according to the following:

* The invoice is divided into a large number of horizontal segments along the y-axis (not shown). Each segment is equivalent to one position in the h-key. If a certain segment includes one or more line elements 15, a mark or tag is placed in the equivalent key position. If not, an empty space, an inverted mark or anything else that differentiates itself from a mark is used instead.

* A v-key for the vertical line elements 14 is generated in a similar manner along the x-axis.

* The h and v-keys are given designations and together constitute a line key. Following this, a search is performed, which means that the current line key is compared with line keys for known invoices 10 that exist in EH's database. This comparison takes into account that individual lines or line elements 14, 15 can vary somewhat in position, plus that the total pattern of lines can be displaced somewhat according to suitable pre-determined values in the x and y directions, horizontally and vertically respectively.

* The line keys in the database are sorted according to the number of markings (tags), which are used to make the searching effective.

* The search results in a pre-determined number of probable candidates for the identity of the current invoice 10. All candidates are associated with a factor of merit or a probability that they are the current invoice 10.

* The identity of the invoice is finally confirmed by carrying out an interpretation of that known as the RCG-value (RCG- ReCoGnition). The RCG-value is a value at a given position that is unique for a certain invoice/supplier or other form. Examples of such

values are bank giro numbers, post-office giro numbers, invoice numbers, total amounts, etc.

The said segments can, for example, form checked patterns that are fine-screened to varying extents according to the relative need for rapid searching.

5 The line keys can even be implemented on objects formed wholly or partly of text and colours. These are assigned line keys from an object area list that includes x and y-keys for the object. The object area list can, for example, consist of positions for certain selected objects. The principles for line maps stated above are even appropriate for objects other than line elements to accomplish identification of forms.

10 If the line keys are not found in the database, this indicates that the invoice is not known, which results in the new line keys being stored in the database that, in this way, is updated in real time.

 If necessary, the operator can, via his computer, manually support the whole or part of the adaptive registration and/or identification of a new form or registered form respectively if
15 several alternative candidates are presented as probable according to the factor of merit.

 In addition, the present invention includes an arrangement for performing the method according to the above.

 The arrangement performs the automatic data acquisition, by means of a means for the same, of forms whose design and information content is not known in advance, by input into
20 the said means together with storage of patterns of the same. It registers in an adaptive manner and learns the design of forms, and includes a computer with the following means for accomplishing the adaptive registration:

 means for generating a form map based on the previously unknown form's design for identifying information contained on the form;

25 means for searching and comparing the form map with stored, recognised maps in a means for storing form maps;

 means for storage of generated form maps in the storage means when they do not coincide with a stored map according to pre-determined limits for agreement;

 means for indicating agreement according to the limits for agreement when
30 agreement is found; and

 means for identification and continued data acquisition of the information content of the form.

The said means are preferably controlled by computer hardware and software, such as, for example:

A scanner for acquisition of data.

An electronic storage medium (hard disk, CD-ROM, etc.) for the means to store
5 information

Signs, icons, signal generators, etc. for indicating purposes.

Filters and comparitors so that the means can search and compare, as well as filters and registers for identification.

On the whole, the means used in the present invention are well known to a skilled
10 person in the technical field, but the way in which they are co-ordinated to achieve the object of the invention is, however, innovative.

In one embodiment of the present invention with reference to Fig. 2, a schematic flow-path is illustrated to show the scanning, identification, interpretation, and validation of a form according to the present invention.

Fig. 2 is divided by dotted lines into partial areas to clarify the different steps in a
15 method according to the invention, whereby the steps constitute the scanning of the form 200, identification of the form 210, interpretation of the form 220, plus validating the form 230.

The form is scanned 200 into EH, and identification 210 follows. Identification consists of generating a line map 212, or alternatively an object area list, whereby an line key is
20 generated. Following this, form 10 is compared 214 with known keys in the form map database, whereby a conformation of identification is obtained via the RCG-value. The next step includes deciding whether the identification was successful 216 according to the conditions "Yes" or "No". If the decision results in "No", a conditional investigation is made to see if there are more candidates in the form of line keys 218. In the answer here is "Yes", a loop in the form of
25 214, 216 and 218 is performed until a successful identification is finally made, or until no further line key candidates are presented 218.

In the case of a successful identification, interpretation 220 of the form then begins by interpreting with the help of the current form map 222, after which validation 230 or
30 evaluation 232 of the fields of the form 10 takes place. As an option, the operator can assist with selection if several alternative fields are found 234.

If the identification 210 is unsuccessful, and no further line keys are presented 218, interpretation 220 is performed in that self-learning with a form definition 224 is accomplished.

The form definition consists of a template or a set of rules that describes the common elements of a specific collection of forms, for example, Swedish invoices. Following this, the RCG-value is interpreted 226 and a decision is made 228 whether the current RCG-value can be found in the form database. If the answer is "Yes", a re-interpretation begins 229, followed by a continued
5 interpretation 222 that leads to validation 232.

If, on the other hand, the answer is "No", validation commences 230, 236, after which the form is saved in the form map database with the line key 238. Prior to steps 236, 238, the operator can, if several field alternatives are found, assist with the self-learning process.

The embodiments of the present invention described above are possible
10 embodiments, but are not intended to limit the invention to such, as further embodiments will be evident to a skilled person in the technical area via the drafts of the enclosed claims.

Claims

1. Method for the automatic data acquisition (200), by means of a means for the same, of forms (10) whose design and information content (19) is not known in advance, by input into the said means, together with storage of patterns of the same, c h a r a c t e r i s e d in that
5 the method is adaptive (224) and includes self-learning and registration of the design of forms, whereby it includes the following steps to accomplish the adaptive registration (238):

generation of a form map (18) based on the previously unknown form's (10) design for identifying (210) information contained on the form;

10 searching and comparing (210, 220) the form map (18) with stored, registered maps in a means for storing form maps;

storage (238) of generated form maps (18) in the storage means when they do not coincide with a stored map according to pre-determined limits for agreement;

indication of agreement according to the limits for agreement when agreement is found; and

15 continued data acquisition (232) for identifying the information content of the form.

2. Method according to claim 1, c h a r a c t e r i s e d in that the form map (18) consists of an object area list with objects (19) contained in the form.

3. Method according to claim 2, c h a r a c t e r i s e d in that the object (19) comprises colours and/or wholly or partly of text.

20 4. Method according to claim 1, c h a r a c t e r i s e d in that the form map (18) constitutes a line map comprising line elements (14,15) from the form (10).

5. Method according to claim 4, c h a r a c t e r i s e d in that the horizontal lines (15) in the line map are used to generate a horizontal key by dividing the form into a pre-determined number of horizontal segments along the y-axis in a cartographic system of co-ordinates, whereby each segment is equivalent to a position in the horizontal key.
25

6. Method according to claims 4 and 5, c h a r a c t e r i s e d in that the vertical lines (14) in the line map are used to generate a vertical key by dividing the form (10) into a pre-determined number of vertical segments along the x-axis in a cartographic system of co-ordinates, whereby each segment is equivalent to a position in the vertical key.

30 7. Method according to claims 5 and 6, c h a r a c t e r i s e d in that at least one line element (14, 15) that is included in a segment is marked in the equivalent key position, and that segments that lack line elements remain unmarked in the equivalent key position.

8. Method according to claims 4-7, characterised in that the horizontal key and/or a vertical key constitute a line key in the line map (18), whereby, during the said searching, the line key generated is compared (214) with line keys stored in the storage means for verifying agreement.

5 9. Method according to claim 8, characterised in that the line keys are sorted in the storage means according to the number of markings.

10 10. Method according to claims 1 and 2, characterised in that the object's horizontal position in the object area list is used to generate a horizontal key by dividing the form into a pre-determined number of horizontal segments along the y-axis in a cartographic system of co-ordinates, whereby each segment is equivalent to a position in the horizontal key.

11. Method according to claims 1 and 2, characterised in that the object's vertical position in the object area list is used to generate a vertical key by dividing the form into a pre-determined number of vertical segments along the x-axis in a cartographic system of co-ordinates, whereby each segment is equivalent to a position in the vertical key.

15 12. Method according to claims 10 and 11, characterised in that at least one object that is included in a segment is marked in the equivalent key position, and that segments that lack objects remain unmarked in the equivalent key position

20 13. Method according to claims 10-12, characterised in that a horizontal key and/or a vertical key constitute an object key in the object area list, whereby, during the said searching, the object key generated is compared with object keys stored in the storage means for verifying agreement.

14. Method according to claim 13, characterised in that the object keys are sorted in the storage means according to the number of markings.

25 15. Method according to claims 1-14, characterised in that the searching (210, 218) results in a pre-defined number of requested probable candidates for the currently searched form.

30 16. Method according to claim 15, characterised in that an operator can support manually (234, 239) the whole or parts of the adaptive registration or identification of the new form or registered forms respectively if several alternative candidates are found as probabilities according to a factor of merit.

17. Method according to claims 1-16, characterised in that the identity of the form is confirmed by the data acquisition of a RCG-value (214).

18. Arrangement for the automatic data acquisition, by means of a means for the same, of forms (10) whose design and information content (19) is not known in advance, by input into the said means together with storage of patterns of the same, characterised in that it learns adaptively and registers the design of forms (10), and includes a computer with the following means for carrying out the adaptive registration (238):

means for generating a form map (18) based on the previously unknown form's (10) design for identifying information contained on the form;

means for searching and comparing the form map with stored, recognised maps in a means for storing form maps;

means for storage of generated form maps (18) in the storage means when they do not coincide with a stored map according to pre-determined limits for agreement;

means for indicating agreement according to the limits for agreement when agreement is found; and

means for identification and continued data acquisition of the information content of the form.

19. Arrangement according to claim 18 characterised in that the form map (18) consists of an object area list with objects (19) contained in the form (10).

20. Arrangement according to claim 19, characterised in that the object (19) comprises colours and/or wholly or partly of text.

21. Arrangement according to claim 18, characterised in that the form map (18) constitutes a line map comprising line elements (14, 15) from the form.

22. Arrangement according to claim 21, characterised in that the horizontal lines (15) in the line map are used to generate a horizontal key by dividing the form into a pre-determined number of horizontal segments along the y-axis in a cartographic system of co-ordinates, whereby each segment is equivalent to a position in the horizontal key.

23. Arrangement according to claims 21 and 22, characterised in that the vertical lines (14) in the line map are used to generate a vertical key by dividing the form into a pre-determined number of vertical segments along the x-axis in a cartographic system of co-ordinates, whereby each segment is equivalent to a position in the vertical key.

24. Arrangement according to claims 22 and 23, characterised in that at least one line element (14, 15) that is included in a segment is marked in the equivalent key

position, and that segments that lack line elements remain unmarked in the equivalent key position.

25. Arrangement according to claims 22-24, characterised in that the horizontal key and/or a vertical key constitute a line key in the line map, whereby, during the said searching, the line key generated is compared with line keys stored in the storage means for verifying agreement.

26. Arrangement according to claim 25, characterised in that the line keys are sorted in the storage means according to the number of markings.

27. Arrangement according to claims 18 and 19, characterised in that the object's (19) horizontal position in the object area list is used to generate a horizontal key by dividing the form into a pre-determined number of horizontal segments along the y-axis in a cartographic system of co-ordinates, whereby each segment is equivalent to a position in the horizontal key.

28. Arrangement according to claims 18 and 19, characterised in that the object's (19) vertical position in the object area list is used to generate a vertical key by dividing the form into a pre-determined number of vertical segments along the x-axis in a cartographic system of co-ordinates, whereby each segment is equivalent to a position in the vertical key.

29. Arrangement according to claims 27 and 28, characterised in that at least one object that is included in a segment is marked in the equivalent key position, and that segments that lack objects remain unmarked in the equivalent key position.

30. Arrangement according to claims 27-29, characterised in that a horizontal key and/or a vertical key constitute an object key in the object area list, whereby, during the said searching, the object key generated is compared with object keys stored in the storage means for verifying agreement.

31. Arrangement according to claim 30, characterised in that the object keys are sorted in the storage means according to the number of markings.

32. Arrangement according to claims 18-31, characterised in that the searching results in a pre-defined number of requested probable candidates for the currently searched form.

33. Arrangement according to claim 32, characterised in that an operator can support manually the whole or parts of the adaptive registration or identification of the new

form or registered forms respectively if several alternative candidates are found as probabilities according to a factor of merit.

34. Arrangement according to claims 18-33, characterised in that the identity of the form is confirmed by the data acquisition of a RCG-value.

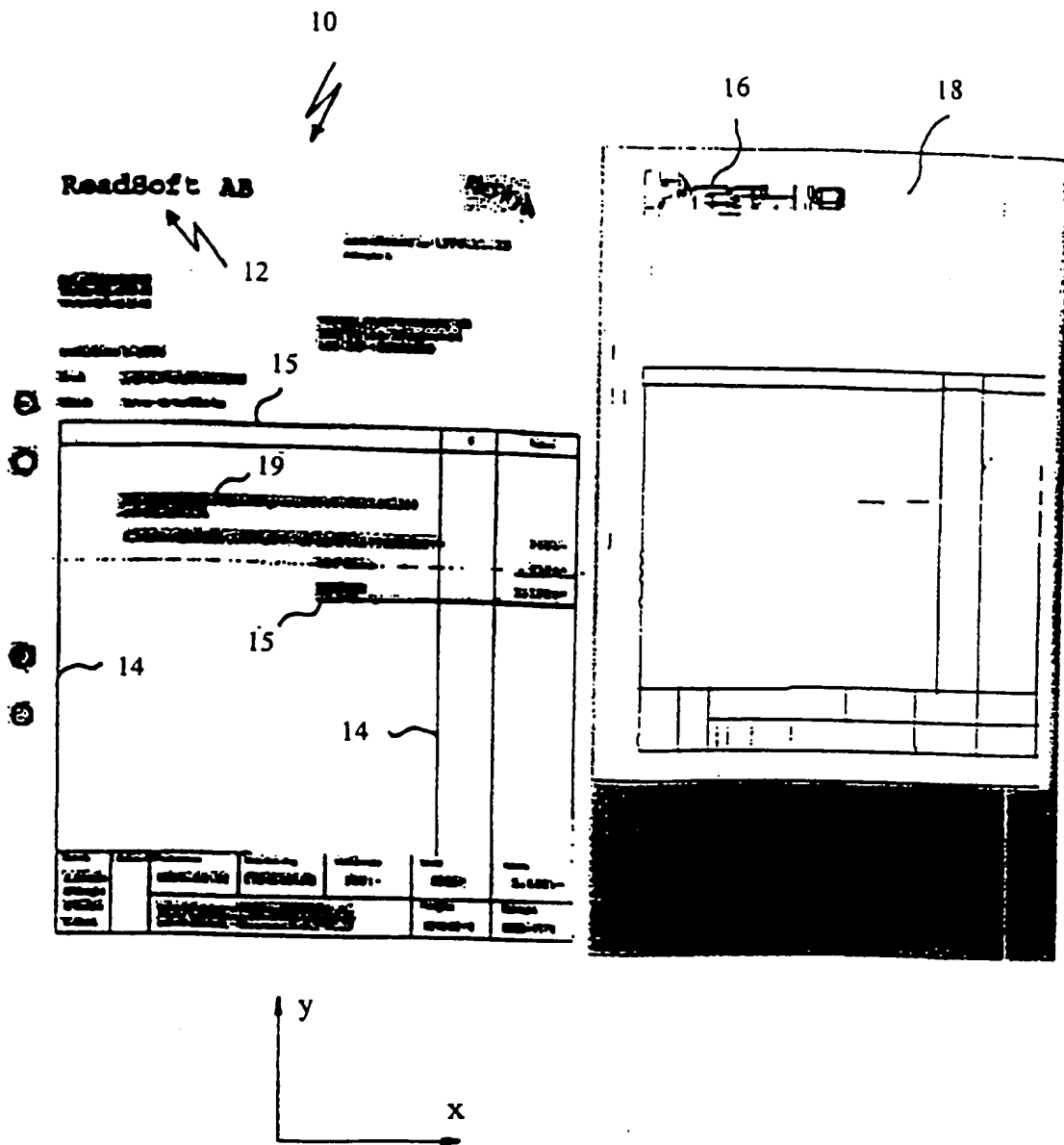


Fig. 1

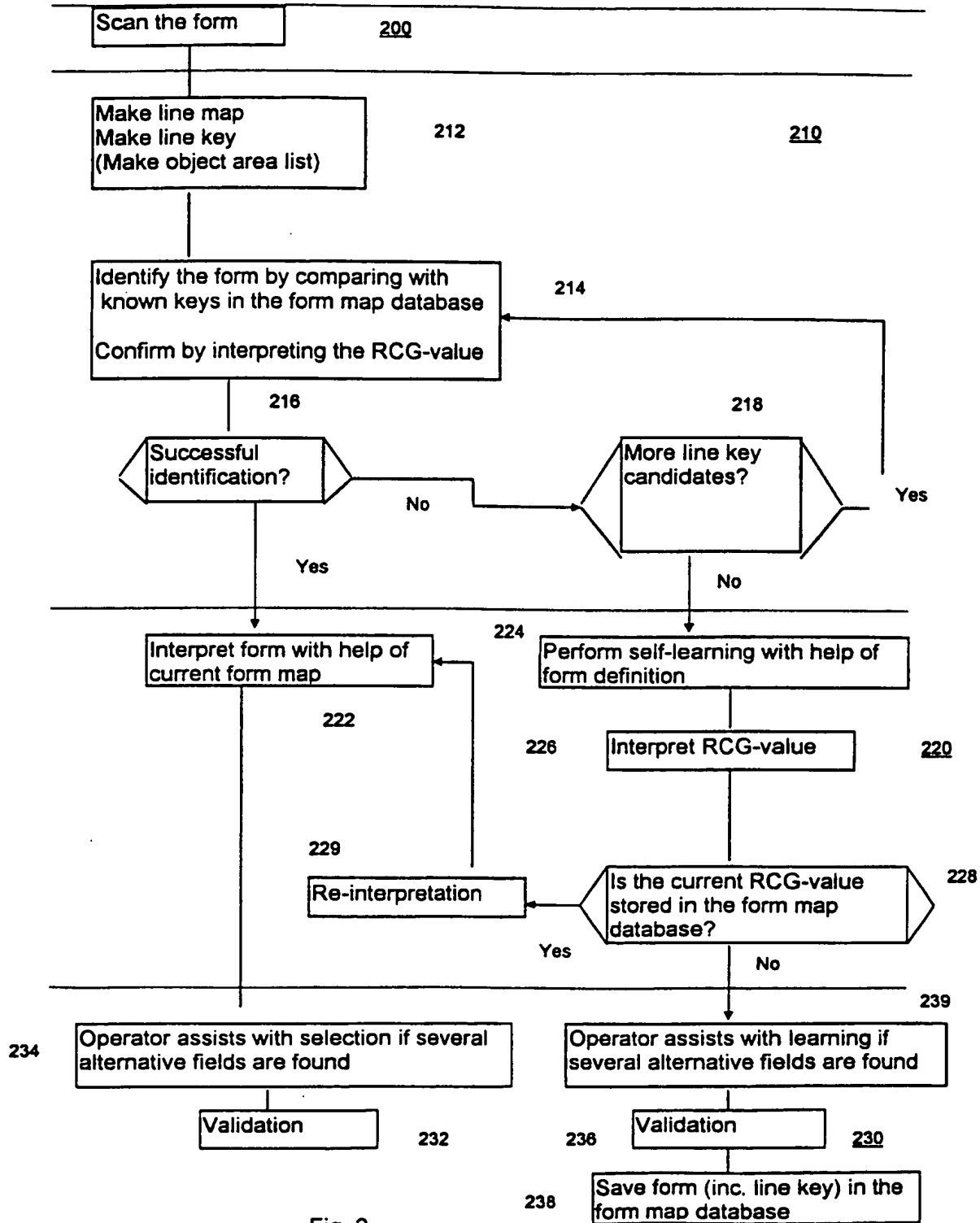


Fig. 2

INTERNATIONAL SEARCH REPORT

International application No.

PCT/SE 98/00602

A. CLASSIFICATION OF SUBJECT MATTER

IPC6: G06K 9/20 // G 06 F 17/24

According to International Patent Classification (IPC) or to both national classification and IPC

B. FIELDS SEARCHED

Minimum documentation searched (classification system followed by classification symbols)

IPC6: G06F, G06K

Documentation searched other than minimum documentation to the extent that such documents are included in the fields searched

SE,DK,FI,NO classes as above

Electronic data base consulted during the international search (name of data base and, where practicable, search terms used)

WPI

C. DOCUMENTS CONSIDERED TO BE RELEVANT

Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
X	US 4933979 A (A. SUZUKI ET AL.), 12 June 1990 (12.06.90), column 2, line 10 - line 51; column 9, line 51 - column 10, line 52 --	1-34
X	US 5140650 A (R.G. CASEY ET AL.), 18 August 1992 (18.08.92), column 2, line 18 - line 65 --	1-34
A	US 5293429 A (A. PIZANO ET AL.), 8 March 1994 (08.03.94), abstract -- -----	1-34

☐ Further documents are listed in the continuation of Box C.
 ☒ See patent family annex.

* Special categories of cited documents:

"A" document defining the general state of the art which is not considered to be of particular relevance

"E" earlier document but published on or after the international filing date

"L" document which may throw doubts on priority claim(s) or which is cited to establish the publication date of another citation or other special reason (as specified)

"O" document referring to an oral disclosure, use, exhibition or other means

"P" document published prior to the international filing date but later than the priority date claimed

"I" later document published after the international filing date or priority date and not in conflict with the application but cited to understand the principle or theory underlying the invention

"X" document of particular relevance: the claimed invention cannot be considered novel or cannot be considered to involve an inventive step when the document is taken alone

"Y" document of particular relevance: the claimed invention cannot be considered to involve an inventive step when the document is combined with one or more other such documents, such combination being obvious to a person skilled in the art

"&" document member of the same patent family

Date of the actual completion of the international search

2 July 1998

Date of mailing of the international search report

06 -07- 1998

 Name and mailing address of the ISA/
 Swedish Patent Office
 Box 5055, S-102 42 STOCKHOLM
 Facsimile No. +46 8 666 02 86

Authorized officer

 Malin Keijser
 Telephone No. +46 8 782 25 00

INTERNATIONAL SEARCH REPORT

Information on patent family members

09/06/98

International application No.

PCT/SE 98/00602

Patent document cited in search report	Publication date	Patent family member(s)	Publication date
US 4933979 A	12/06/90	JP 63155386 A	28/06/88
US 5140650 A	18/08/92	CA 2000012 A,C	02/08/90
		JP 2231692 A	13/09/90
US 5293429 A	08/03/94	JP 5217019 A	27/08/93